

Parameter-tuning Networks: Experiments and Active Walk Model

Xiao-Pu Han^{1,c,*}, Chun-Dong Hu², Zhi-Min Liu², and Bing-Hong Wang^{1,3†}

¹ *Department of Modern Physics, University of Science and Technology of China, Hefei, 230026, China*

² *Institute of Plasma Physics, Chinese Academy of Sciences, Hefei, 230031, China*

³ *Shanghai Academy of System Science and University of Shanghai for Science and Technology, Shanghai, 200093 China*

(Dated: February 28, 2008)

The tuning process of a large apparatus of many components could be represented and quantified by constructing parameter-tuning networks. The experimental tuning of the ion source of the neutral beam injector of HT-7 Tokamak is presented as an example. Stretched-exponential cumulative degree distributions are found in the parameter-tuning networks. An active walk model with eight walkers is constructed. Each active walker is a particle moving with friction in an energy landscape; the landscape is modified by the collective action of all the walkers. Numerical simulations show that the parameter-tuning networks generated by the model also give stretched exponential functions, in good agreement with experiments. Our methods provide a new way and a new insight to understand the action of humans in the parameter-tuning of experimental processes, is helpful for experimental research and other optimization problems.

PACS numbers: 89.75.-k, 89.75.Hc, 05.40.Fb

Large apparatuses pay more and more important role in the experimental research of modern physics, such as the high-energy accelerators and controlling nuclear fusion devices. A large apparatus made up of many components, each component needs be tested and tuned separately and then collectively, before experiments using the apparatus as a whole can be conducted. And there are a large number of parameters to be tuned. The parameter tuning processes are often complicated and onerous, strongly depend on the the experience of experimenter. How can we model the adjustment process of these parameters and make some sense out of it?

For example, in a Tokamak in plasma physics, in the ion source segment of the neutral-beam injector system alone, there are eight major parameters to be adjusted. The experimenter sets the parameter values, turns on the equipment, and measures the outcome of a certain quantity Q . If the Q obtained does not meet the mark Q_0 , say, the whole process is repeated, with a new set of parameters. The adjustment process stops when Q is equal or very close to Q_0 . The simultaneous adjustment of a large number of parameters is a complicated process, which is based on the feedback from previous Q values obtained, the experimenter's experience, and the limitation of the hardware that control the parameters.

To quantify this complicated process, a parameter-tuning network can be constructed as follows. Assume that N parameters $u_i (i = 1, 2, \dots, N)$ are involved. The experimenter's serial adjustment of $\mathbf{u} \equiv (u_1, \dots, u_N)$ can be represented by a sequence of connected dots in the N -dimensional \mathbf{u} space. For each choice of \mathbf{u} , Q is measured, so that $Q = Q(\mathbf{u})$ but the functional form is

unknown due to the complexity of the equipment, which is like a black box.

The sequence of \mathbf{u} dots is then projected onto the (u_1, u_2) plane, say. The dots on the same vertical line along the u_3 direction are allowed to collapse to one point, called a node. A line connecting two nodes is called an edge [1, 2, 3]. Since in real situations, the parameters selected by the experimenter in different adjustments may partially overlap with each other, there may be more than one edge connecting two nodes. These edges are directed. For simplicity, we make the approximation of collapsing all the edges between two nodes that have the same direction as one single edge with the same direction. Consequently, between two nodes, there are at most two edges with opposite directions, resulting in a *directed* parameter-tuning network. A *non-directed* network is formed from the directed one, by collapsing all the edges between any two nodes and removing the directions.

In a parameter-tuning network, there are N_n nodes, say. For a directed parameter-tuning network node $j (j = 1, 2, \dots, N_n)$, let $N_+(j)$ be the number of in-going edges, $N_-(j)$ the number of out-going edge.

Now count the number of N_+ that has the value k , and call it $p_+(k)$; similarly for $p_-(k)$. The corresponding function in the non-directed network case is denoted by $p(k)$. k is called the "degree" of a node. $p_+(k)$ is called the "in-degree distribution", $p_-(k)$ the "out-degree distribution", and $p(k)$ the "degree distribution." Next define the corresponding "cumulative degree distributions" $P_+(k)$, $P_-(k)$ and $P(k)$ by

$$P_+(k) \equiv \sum_{k'=k}^{k_m} p_+(k') \quad (1)$$

*Electronic address: hxp@mail.ustc.edu.cn

†Electronic address: bhwang@ustc.edu.cn

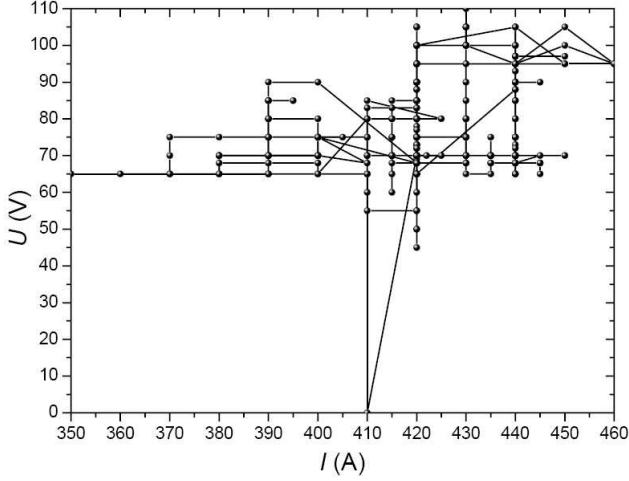


FIG. 1: The non-directed parameter network on the (U, I) plane derived from the experimental data, where U is the working voltage of the cathode gas valve and I the filament current.

$$P_{-}(k) \equiv \sum_{k'=k}^{k_m} p_{-}(k') \quad (2)$$

$$P(k) \equiv \sum_{k'=k}^{k_m} p(k') \quad (3)$$

Here k_m is the maximum k that the p function is non-zero, and all of the distribution functions are not unitary. Note that by definition, the P 's are monotonic decreasing functions, while the p 's may not be monotonic at all. The reason for introducing the P 's is that we want to fit them to monotonic decreasing functions such as stretched exponents or power laws.

In our experiments studying the ion source of neutral beam injector system for HT-7 Tokamak, eight control parameters are involved ($N = 8$). These include the filament current, magnet current, arc voltage, cathode gas valve voltage, anode gas valve voltage, etc. Each of the parameters has 10 to 30 discrete, adjustable setting values. The aim of the experiment is to find out which set of parameters will give a strong and stable discharge, measured by the arc current intensity Q . Each parameter is set by a dial which can be turned left or right between two extreme positions. When the extreme position is reached, the experimenter has to turn the dial back, reversing the direction of turning.

From our experimental data, the sequence of parameters in the 8-dimensional \mathbf{u} space is first generated and the corresponding directed and non-directed tuning-parameter networks are constructed. The non-directed network is shown in FIG. 1. The P 's are obtained, and fitted nicely with stretched exponential functions (FIG. 2) such that

$$P(k) \sim \exp(-ak^\gamma) \quad (4)$$

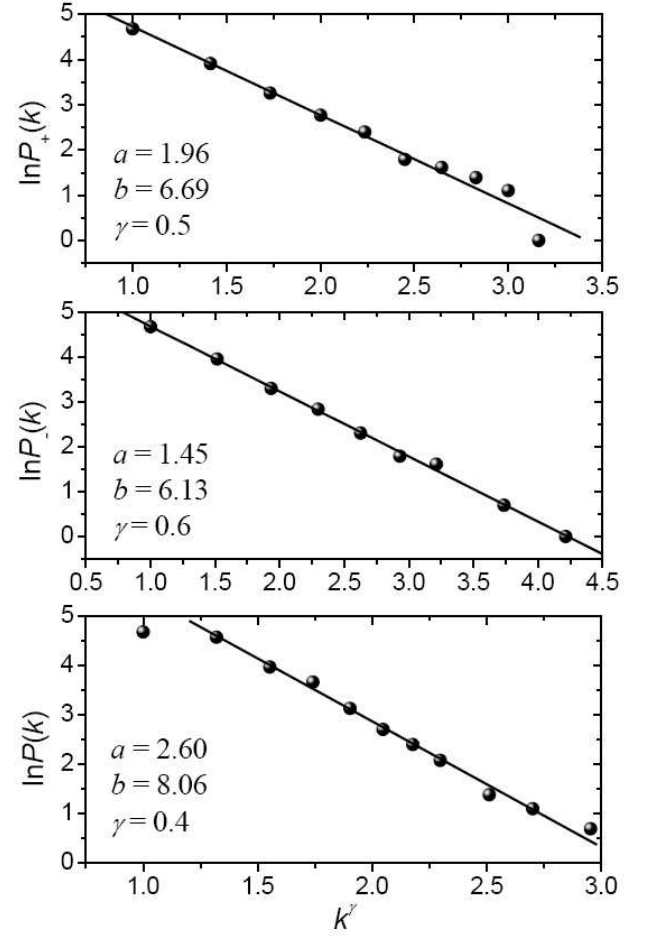


FIG. 2: Experimental curves of $P_{+}(k)$, $P_{-}(k)$ and $P(k)$, fitted to stretched exponential functions. The corresponding (directed and non-directed) parameter-tuning networks exist on the (U, I) plane, where U is the working voltage of the cathode gas valve and I the filament current. Its non-directed structure has been shown in FIG. 1. The experimental data obtained from the tuning of the ion source in a whole experiment season, the sequence includes 789 parameter sets.

or, equivalently,

$$\ln P(k) = -ak^\gamma + b \quad (5)$$

and similarly for $P_{+}(k)$ and $P_{-}(k)$. Such property is also found in the parameter-tuning networks derived from the data of the tuning experiments in other experiment season of the ion source (each of them includes 600 to 800 parameter sets).

The experimental adjustment of the parameters is somehow correlated by the experimenter, and the process can be modeled by an active walk (AW) model [4]. In our AW model here, $N = 8$; each u_i has adjustable values of $1, 2, \dots, 20$, say. The adjustment of the i th parameter is represented by the movement of an active walker on a 1D landscape potential $V_i(x_i)$. The allowable x_i for the i th walker are the integers $(1, 2, 3, \dots, 100)$, the same for all

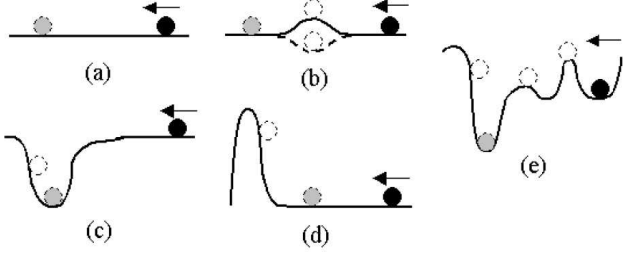


FIG. 3: Sketch of possible subwalks of a particle. The solid dot, gray dot and open circle represent the initial position, final position and some intermediate position of the particle, respectively; the arrow indicates the initial direction of the particle. (a) The particle moves and stops on a plateau. (b) The particle gets over a low potential barrier or a shallow well. (c) The particle stops in a deep well. (d) The particle rebounds on a high potential barrier. (e) The particle moves on a general landscape, and ends up in a deep well after many steps.

i. These 100 numbers are mapped to the u_i , such that if the i th walker ends in the region $[1, 5]$ on the x_i axis, u_i will assume the value of 1. Similarly, the region of $[6, 10]$ is mapped to $u_i = 2$, etc. This mapping can be attained by

$$u_i(t) = \text{Int}[(X_i(t) - 1)/5] + 1 \quad (6)$$

where $X_i(t)$ is the position of the i th walker on the x_i axis at time t , and Int means taking the integral part of the number. This mapping of x_i to u_i has the effect of making two consecutive sets of adjusted parameters more likely to partially overlap with each other, and ensures the occurrence of smooth V_i vs x_i curves.

In the simulations, all V_i start flat at time $t = 0$, and are updated simultaneously at each time step $t (= 1, 2, \dots)$, according to a rule to be specified below. But between two consecutive time steps, each walker moves a few steps in a subwalk in its own V_i , independent of other walkers. In a subwalk, the walker does not change V_i . The subwalk is like a particle rolling on a landscape with friction, with the following rules (with the subscript i removed for the sake of clarity). The subwalk time is labeled by $\tau (= 0, 1, 2, \dots)$.

(i). At time $t = 0$, the particle is arbitrarily placed on the x axis.

(ii). At time t , the particle is given energy K_0 at $\tau = 0$. (K_0 is a parameter fixed in the model.)

(iii). At subwalk time τ , the particle moves left or right with equal probability. However, after each move, the particle loses energy ε , and the energy difference $V(x(\tau)) - V(x(\tau + 1))$ which could be positive or negative is added to its energy. Consequently, at time τ , the particle already moves τ steps, and its energy becomes

$$K(\tau) = K_0 + V_0 - V(x(\tau)) - \tau\varepsilon \quad (7)$$

where $V_0 \equiv V(x(0))$, the potential at the initial position of the particle at $\tau = 0$.

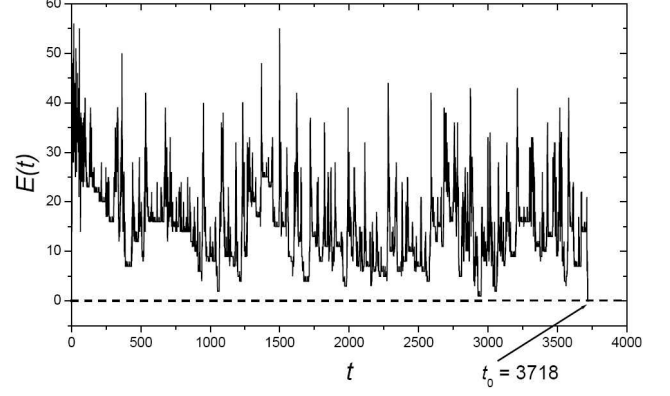


FIG. 4: Dependence of the error function on time, from the eight-parameter AW model simulation.

(iv). The particle can get over a potential barrier in V that is lower than K , and can rebound if the barrier is higher than K .

(v). When the particle reaches the left boundary ($x = 1$) or the right boundary ($x = 100$), it reverses direction and continues its subwalk.

(vi). The particle stops when it walks into a potential well and cannot get out with its energy K , or exhausts its energy on a plateau.

These possibilities are sketched in FIG. 3. The reflecting boundary condition in item (v) corresponds to the real experimental situation where a dial with a limited range is used.

After all N particles stop, the time clock increases from t to $t + 1$; the particle's position at the end of its subwalk is taken to be $X(t + 1)$, which is the starting position of the subwalk at time $t + 1$. (The subwalks of the particles may not stop after the same number of subwalk steps; those stop first will sit there and wait for the last particle to stop.) The landscape $V_i(x_i)$ is updated by the landscaping rule:

$$V_i(x_i; t + 1) = \begin{cases} V_i(x_i; t) + W(x_i - X_i(t + 1)), & E(t) \geq E_a(t); \\ V_i(x_i; t) - W(x_i - X_i(t + 1)), & E(t) < E_a(t). \end{cases} \quad (8)$$

where $t = 0, 1, 2, \dots$. In our numerical simulations below, the (i -independent) landscaping function W is given by $W(0) = 4$, $W(\pm 1) = 3$, $W(\pm 2) = 1$, and $W = 0$ otherwise. In Eq. (8), the error function $E(t)$ is assumed to be

$$E(t) = |Q(\mathbf{u}(t)) - Q_0| \quad (9)$$

where $\mathbf{u}(t)$ is obtained from Eq. (6); $E_a(t)$ is the average error counting the last n time steps, given by

$$E_a(t) = \frac{1}{n} \sum_{t'=t-n+1}^t E(t') \quad (10)$$

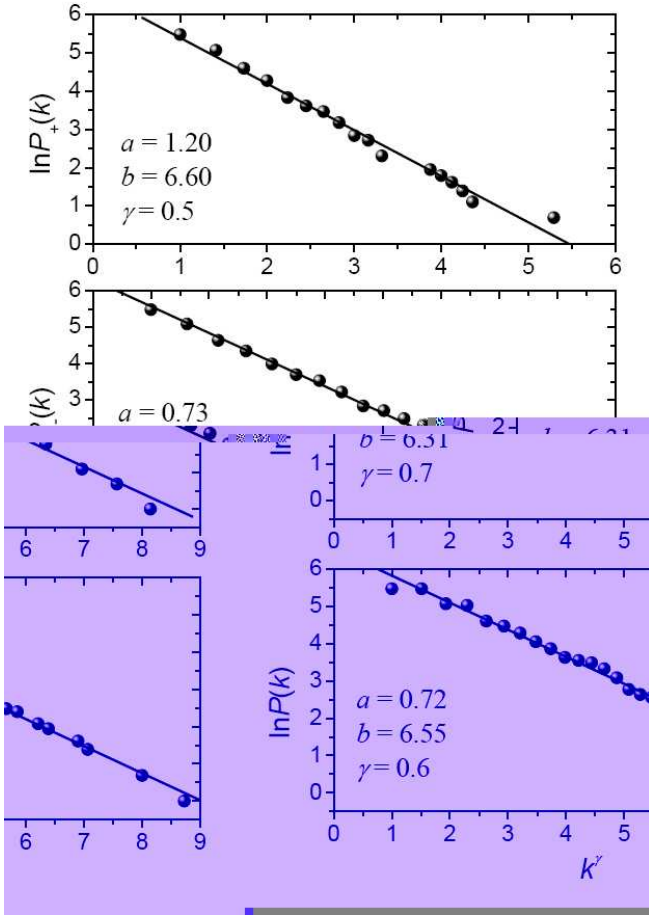


FIG. 5: Numerical curves of $P_+(k)$, $P_-(k)$ and $P(k)$ obtained from the parameter-tuning networks from a eight-parameter AW model simulation, fitted to stretched exponent functions.

In computer simulations, we pretend that we know Q_0 and the $Q(\mathbf{u})$ function. In reality, the former is known to the experimenter; the latter can be roughly inferred from experimental data. Note that while the subwalks of the particles are independent of each other, the updating of their landscapes is affected by their collective effort through $Q(\mathbf{u})$ in $E(t)$.

The eight-parameter AW model is simulated with t going from 0, and the simulation is stopped when the optimal parameter set is gotten. The function Q is assumed

to be

$$Q(\mathbf{u}) = |u_1 - 10| + |u_2 - 10| + \cdots + |u_8 - 10| \quad (11)$$

The parameters used are: $K_0 = 23$, $\varepsilon = 1$, $n = 5$ and $Q_0 = 0$. This means that $\mathbf{u} = (10, 10, \dots, 10)$ is the one and only optimal parameter set.

The error function $E(t)$ is displayed in FIG. 4, which shows that the optimal \mathbf{u} is first obtained at $t_0 = 3718$. The 700 $\mathbf{u}(t)$ dots (with $300 < t \leq 1000$, such number of $\mathbf{u}(t)$ is similar with the experimental data, and the initialization process when $t < 300$ is ignored) in the \mathbf{u} space are projected onto the (u_1, u_2) plane to obtain the directed and non-directed networks. The corresponding cumulative degree distributions are given in FIG. 5. Good stretched-exponential fits are obtained, in agreement with the experimental findings in FIG. 2.

In the AW model, the subwalks kind of mimicking the action of the experimenter—we think that is why the AW model and the experiments give similar P functions. Human dynamics [5, 6] have been studied in other situations by different models. Our results imply that our network-based statistics and the AW model is helpful in the understanding of human action in the experimental processes and very likely in many other optimization projects. The stretched exponent property of the parameter-tuning network imply the human action in the optimization process could be having some quantitative principles, which is important for the experimental studies. Stretched exponent distributions or relations exist widely in many other systems [7, 8, 9, 10]. Since AW-like dynamics is very common in nature [4], our model may reveal a new mechanism of stretched exponent distribution. By modifying the subwalks and tuning their parameters one may shorten t_0 , the shortest time to find the optimal parameters, which would be of interest to the experimental researchers. Finally, the AW model mimics an optimization process in real experiment, could be useful in the optimization projects in many other systems.

We thank Lui Lam, Pin-Qun Jiang, Da-Ren He and Tao Zhou for useful discussions, and Sheng Liu, Shi-Hua Song, Jun Li and Yuan-Lai Xie for experimental help. This work is partially supported by China 973 Program (Grant No. 2006CB705500) and NSF China (Grant Nos. 60744003, 10635040, 10532060, 10472116, and 10575105).

-
- [1] D. J. Watts, S. H. Strogatz, *Nature* **393**, 440 (1998).
 - [2] A. -L. Barabási, R. Albert, *Science* **286**, 509 (1999).
 - [3] R. Albert, and A. -L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [4] L. Lam, *Int. J. Bifurcation & Chaos* **15**, 2317 (2005); *Int. J. Bifurcation & Chaos* **16**, 239 (2006).
 - [5] A. -L. Barabási, *Nature* **435**, 207 (2005).
 - [6] A. Vázquez, *Phys. Rev. Lett.* **95**, 248701 (2005).
 - [7] R. Xulvi-Brunet, and I. M. Sokolov, *Phys. Rev. E* **66**,

- 026118 (2002).
- [8] A. J. Holanda, I. T. Pisa, O. Kinouchi, A. S. Martinez, and E. E. S. Ruiz, *Physica A* **344**, 530 (2004).
- [9] J. Laherrère and D. Sornette, *Eur. Phys. J. B* **2**, 525 (1998).
- [10] B. Sturman, E. Podivilov, and M. Gorkunov, *Phys. Rev. Lett.* **91**, 176602 (2005).